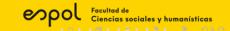
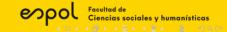
# Vectores Autoregresivos - VAR

Leonardo Sánchez Aragón

August 13, 2020



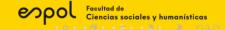
- Cuando no tenemos la certeza de que una variable es exógena, una extensión del ande función de trasnferencia es VAR.
- Consideremos el caso de dos variables que serán endógenas: la secuencia de valores  $\{y_t\}$  son efectados por valores pasados de si mismo y los valores actuales y pasados de  $\{z_t\}$ . De igual fomar sucede con la secuencia  $\{z_t\}$   $y_t = b_{10} b_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \varepsilon_{yt}$   $z_t = b_{20} b_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \varepsilon_{zt}$
- ► Este es un Vector Autoregresivo de orden 1 VAR(1). Este sistema se conoce como VAR Estructural.
- Se asume que: (i)  $y_t$  y  $z_t$  son estacionarias; (ii)  $\varepsilon_{yt}$  y  $\varepsilon_{zt}$  son ruido blanco con desviaciones  $\sigma_v$  y  $\sigma_z$ ; y (iii)  $\varepsilon_{vt}$  y  $\varepsilon_{zt}$  están no correlacionados.



- ▶ El coeficiente  $b_{12}$  es el efecto contemporáneo de  $z_t$  sobre  $y_t$ , y  $\gamma_{12}$  es el efecto de  $z_{t-1}$  sobre  $y_t$ .
- Si  $b_{21} \neq 0$ , entonces  $\varepsilon_{y_t}$  tiene un efecto indirecto contemparaneo sobre  $z_t$ , y si  $b_{12} \neq 0$ , entonces  $\varepsilon_{z_t}$  tiene un efecto indirecto contemparaneo sobre  $y_t$ .
- Las ecuaciones previas no se pueden estimar usando MCO ya que  $y_t$  tiene un efecto contemporaneo sobre  $z_t$  y esta sobre  $y_t$ . Las estimaciones de MCO sufrirían un sesgo de ecuación simultánea ya que los regresores y los términos de error estarían correlacionados.
- ▶ Reordenando el sistema, se pudiera estimar. Matricialmente tenemos,

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t$$



▶ Donde las matrices de  $Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t$ , son

$$B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}, x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}, \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}, \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \varepsilon_t = \begin{bmatrix} \varepsilon_{y_t} \\ \varepsilon_{z_t} \end{bmatrix}$$

▶ Premultiplicando por  $B^{-1}$ , tenemos

$$x_t = A_0 + A_1 x_{t-1} + e_t$$

Donde 
$$A_0 = B^{-1}\Gamma_0$$
,  $A_1 = B^{-1}\Gamma_1$  y  $e_t = B^{-1}\varepsilon_t$ 

▶ Se define  $a_{i0}$  como el i-esimo elemento del vector  $A_0$ ,  $a_{ij}$  es el elmento de la fila i y columna j de la matriz  $A_1$ ,  $e_{it}$  es el elemento i del vector  $e_t$ .

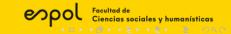


► Entonces, el sistema se puede reescribir como

$$y_t = a_{10} + a_{11}y_{t-1} + a_{12}z_{t-1} + e_{1t}$$
  

$$z_t = a_{20} + a_{21}y_{t-1} + a_{22}z_{t-1} + e_{2t}$$

- Esta forma de representar el VAR se denomina forma estandar.
- Cabe destacar que los errores  $e_t = B^{-1}\varepsilon_t$   $e_{1t} = \left(\varepsilon_{y_t} b_{12}\varepsilon_{z_t}\right) / \left(1 b_{12}b_{21}\right)$   $e_{2t} = \left(\varepsilon_{z_t} b_{21}\varepsilon_{y_t}\right) / \left(1 b_{12}b_{21}\right)$
- Dado que  $\varepsilon_{z_t}$  y  $\varepsilon_{y_t}$  son ruidos blancos, entonces  $e_{1t}$  y  $e_{2t}$  tiene media cero, varianza constante y son individualmente no correlacionados.



El valor esperado es:

$$Ee_{1t} = E\left(\varepsilon_{yt} - b_{12}\varepsilon_{zt}\right)/\left(1 - b_{12}b_{21}\right) = 0$$

La varianza es:

$$Ee_{1t}^{2} = E \left[ \left( \varepsilon_{yt} - b_{12} \varepsilon_{zt} \right) / \left( 1 - b_{12} b_{21} \right) \right]^{2}$$
$$= \left( \sigma_{y}^{2} + b_{12}^{2} \sigma_{z}^{2} \right) / \left( 1 - b_{12} b_{21} \right)^{2}$$

La correlación es

$$Ee_{1t}e_{2t} = E \left[ \left( \varepsilon_{yt} - b_{12}\varepsilon_{zt} \right) \left( \varepsilon_{zt} - b_{21}\varepsilon_{yt} \right) \right] / \left( 1 - b_{12}b_{21} \right)^2$$
  
=  $- \left( b_{21}\sigma_y^2 + b_{12}\sigma_z^2 \right) / \left( 1 - b_{12}b_{21} \right)^2$ 

▶ Si  $b_{12} = b_{21} = 0$  los errores serán no correlacionados.



Es útil escribir la matriz de varianza y covarianza

$$\Sigma = \left[ egin{array}{cc} \mathsf{var}\left(e_{1t}
ight) & \mathsf{cov}\left(e_{1t},e_{2t}
ight) \ \mathsf{cov}\left(e_{1t},e_{2t}
ight) \end{array} 
ight]$$

► En forma compacta tenemos

$$\Sigma = \left[ egin{array}{ccc} \sigma_1^2 & \sigma_{12} \ \sigma_{21} & \sigma_2^2 \end{array} 
ight]$$

# Estacionariedad y Estabilidad

La ecuación  $x_t = A_0 + A_1 x_{t-1} + e_t$ , puede resolverse de manera recursiva como

$$x_t = (I + A_1 + ... + A_1^n)A_0 + \sum_{i=0}^n A_1^i e_{t-i} + A_1^{n+1} x_{t-n-1}$$

- Condición de estabiliadad: Se requiere que la expresión  $A_1^n$  se desvanezca a medida que n tiende al infinito.
- ▶ Dicho de otra forma: se requiere que las raíces de  $(1 a_{11}L)(1 a_{22}L) (a_{12}a_{21}L^2)$  estén fuera del círculo unitario.

# Estacionariedad y Estabilidad

ightharpoonup Si la condición de estabilidad se cumple la solución particular para  $x_t$  es

$$x_t = \mu + \sum_{i=0}^{\infty} A_1^i e_{t-i}$$

$$egin{aligned} ar{y} &= \left[ a_{10} \left( 1 - a_{22} 
ight) + a_{12} a_{20} 
ight] / \Delta \ ar{z} &= \left[ a_{20} \left( 1 - a_{11} 
ight) + a_{21} a_{10} 
ight] / \Delta \ \Delta &= \left( 1 - a_{11} 
ight) \left( 1 - a_{22} 
ight) - a_{12} a_{21} \end{aligned}$$

# Estacionariedad y Estabilidad

- ▶ Si tomamos el valor esperado de  $x_t = \mu + \sum_{i=0}^{\infty} A_1^i e_{t-i}$ , el valor esperado incondicional es  $\mu$
- La matriz de varianza y covarianza es

$$E(x_t - \mu)^2 = E\left[\sum_{i=0}^{\infty} A_1^i e_{t-i}\right]^2$$

$$= (I + A_1^2 + A_1^4 + A_1^6 + \cdots) \Sigma$$

$$= [I - A_1^2]^{-1} \Sigma$$

Donde

$$Ee_t^2 = \left[ egin{array}{c} e_{1t} \ e_{2t} \end{array} \right] \left[ egin{array}{c} e_{1t} \ e_{2t} \end{array} \right] = \Sigma$$

#### Estimación

► Considere el siguiente modelo

$$x_t = A_0 + A_1 x_{t-1} + A_2 x_{t-2} + \dots + A_p x_{t-p} + e_t$$

- Se asume que las variables incorporadas en el VAR responde a un modelo económico (Sims, 1980).
- Este sistema de ecuaciones se estima por Minimo Cuadrados Generalizados.
- Los coeficientes estimados son idénticos a los obtenidos al estimar ecuación por ecuación, ya que todas las regresiones tienen variables idénticas en el lado derecho (Zellner, 1962), pero las desvaciones standard pueden diferir si las ecuaciones están relacionadas
- Otro método de estimacion es usar Máxima Verosimilitud, pero se asume que se conoce la distribución de los errores del sistema.



### Estimación

- ▶ La matriz  $A_0$  contiene n parámetros, y cada matriz  $A_i$  contiene  $n^2$  parámetros; por lo tanto, se necesitan estimar los coeficientes  $n + pn^2$ .
- ► El VAR estará sobreparamizado porque muchas de estas estimaciones de coeficientes serán insignificantes.
- ➤ Solución: imponer restricciones sobre los parámetros del VAR estructural. Pero la imposición incorrecta de restricciones cero puede desperdiciar información importante. Además, es probable que los regresores sean altamente colineales, por lo que las pruebas t de coeficientes individuales no son guías confiables para reducir el modelo.

- Cuando estima el modelo en forma estandar hay 6 parámetros estimados  $(a_{10}, a_{20}, a_{11}, a_{12}, a_{21}, a_{22})$  y los valores de  $var(e_{1t})$ ,  $var(e_{2t})$ ,  $cov(e_{1t}, e_{2t})$
- ► El VAR estructural tiene 10 paráametros:  $(b_{10}, b_{20}, \gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22})$ , los efectos contemporáneos  $(b_{12} \text{ y } b_{21})$  y los valores de  $\sigma_y^2$ ,  $\sigma_z^2$
- Para recuperar los parámetros originales debemos imponer una restricción al modelo estructural.

Una poropuesta es la sugerida por Sims 1980. Si estamos dispuesto a eliminar el impacto de  $y_t$  sobre  $z_t$  (es decir,  $b_{21} = 0$ ), el sistema se hace asimétrico. Esto permite recuperar todos los coeficientes.

$$e_{1t} = \varepsilon_{yt} - b_{12}\varepsilon_{zt}$$

$$e_{2t} = \varepsilon_{zt}$$

Tal que

$$ext{var}\left(e_1
ight)=\sigma_y^2+b_{12}^2\sigma_z^2 \ ext{var}\left(e_2
ight)=\sigma_z^2 \ ext{cov}\left(e_1,e_2
ight)=-b_{12}\sigma_z^2 \ ext{}$$

▶ Viendo la restricción desde otra perspectiva, tenemos.

$$\underbrace{\left[\begin{array}{cc} 1 & b_{12} \\ 0 & 1 \end{array}\right]}_{\mathcal{C}} \left[\begin{array}{c} y_t \\ z_t \end{array}\right] = \left[\begin{array}{c} b_{10} \\ b_{20} \end{array}\right] + \left[\begin{array}{cc} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{array}\right] \left[\begin{array}{c} y_{t-1} \\ z_{t-1} \end{array}\right] + \left[\begin{array}{c} \varepsilon_{yt} \\ \varepsilon_{zt} \end{array}\right]$$

▶ Pre multiplicando por  $B^{-1}$ , tenemos

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} - b_{12}b_{20} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} - b_{12}\gamma_{21} & \gamma_{12} - b_{12}\gamma_{22} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} - b_{12}\varepsilon_{zt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$y_t = a_{10} + a_{11}y_{t-1} + a_{12}z_{t-1} + e_{1t}$$
  
 $z_t = a_{20} + a_{21}y_{t-1} + a_{22}z_{t-1} + e_{2t}$ 

- El supuesto  $b_{21}=0$  significa que  $y_t$  no tiene un efecto contemporáneo sobre  $z_t$ . La restricción se manifiesta de tal manera que los shocks  $\varepsilon_{y_t}$  y  $\varepsilon_{z_t}$  afectan el valor contemporáneo de  $y_t$ , pero solo los choques  $\varepsilon_{z_t}$  afectan el valor contemporáneo de  $z_t$ . Los valores observados de  $e_{2t}$  se atribuyen completamente a shock puros a la secuencia  $z_t$ . Esta descomposición triangular de los residuos se llama descomposición de Choleski.
- De hecho, el resultado es bastante general. En un VAR de n variables, B es una matriz  $n \times n$  ya que hay n residuos de regresión y n shochs estructurales. La identificación exacta requiere que se coloquen  $(n^2 n)/2$  restrictions entre los residuos de regresión. Dado que la descomposición de Choleski es triangular, se fuerza exactamente a  $(n^2 n)/2$  valores de la matriz B a ser igual a cero.

- ▶ Hay un problema si las variables en un VAR no son estacionarias? Sims (1980) y Sims, Stock y Watson (1990) recomendaron no diferenciar, incluso si las variables contienen una raíz unitaria.
- Argumentan que el objetivo de un análisis VAR es determinar las interrelaciones entre las variables, no determinar las estimaciones de los parámetros.
- ▶ El argumento principal en contra de la diferenciación es que "desecha" información sobre los comovimientos en los datos (como la posibilidad de relaciones de cointegración).
- Del mismo modo, se argumenta que no es necesario detrend los datos. En un VAR, una variable de tendencia estará bien aproximada por una raíz unitaria más drift. Sin embargo, la opinión mayoritaria es que la forma de las variables en el VAR debe imitar el verdadero proceso de generación de datos.



## Selección de rezagos

- Cuántos rezagos incluir en el estimación?
- ▶ Un enfoque es estimar muchos modelos VAR con diferentes números de rezagos y comparar su ajuste.
- La literatura utiliza un conjunto de criterios de selección: el criterio de información de Akaike (AIC), el criterio de información bayesiano de Schwarz (SBIC) y el criterio de información de Hannan y Quinn (HQIC).

$$AIC = T \ln |\Sigma| + 2N$$
  
SBC =  $T \ln |\Sigma| + N \ln(T)$ 

Donde  $|\Sigma|$  es el determinante de la matriz de varianza y covarianza de los residuos, y N es el número total de parámetros estimados en todas las ecuaciones. Agregar regresores adicionales reducirá  $\ln |\Sigma|$  a expensas de aumentar N.

► El modelo VAR tiene un representaciobn VMA.

$$x_t = \mu + \sum_{i=0}^{\infty} A_1^i e_{t-i}$$

- La representacion VMA permite expresar las vairanles  $y_t$  y  $z_t$  en terminos de los errores  $(e_{1t}, e_{2t})$  presents y pasados
- ► Esta representacion permite delinear a traves del tiempo el efecto de los shock sobre las variabels presentes en el VAR

► A partir de la representación VMA, tenemos

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t-i} \\ e_{2t-i} \end{bmatrix}$$

▶ Recordando que  $e_t = B^{-1}\varepsilon_t$ , es decir,

$$\left[\begin{array}{c}e_{1t}\\e_{2t}\end{array}\right] = \frac{1}{1 - b_{12}b_{21}} \left[\begin{array}{cc}1 & -b_{12}\\-b_{21} & 1\end{array}\right] \left[\begin{array}{c}\varepsilon_{yt}\\\varepsilon_{zt}\end{array}\right]$$

▶ Podemos unir las expresiones previias y obtener

$$\begin{bmatrix} y_t \\ z_i \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \frac{1}{1 - b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix}$$

La expresión previa puede ser resumida mediante la simplificación de la matriz  $2 \times 2$ 

$$\phi_i = rac{\mathcal{A}_1^i}{1 - b_{12}b_{21}} \left[ egin{array}{cc} 1 & -b_{12} \ -b_{21} & 1 \end{array} 
ight]$$

lacktriangle Esto permite reexpresar en modelo en función de los valores de  $arepsilon_t$ 

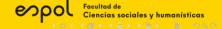
$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix}$$

$$x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$$



- Los elementos  $\phi_{jk}(0)$  son multiplicadores de impacto. Si  $\phi_{12}(0)$  es el instantáneo impacto de un cambio de una unidad de  $\varepsilon_{z_t}$  en  $y_t$ .
- Los elementos  $\phi_{11}(1)$  y  $\phi_{12}(1)$  son las respuestas luego de un período de los cambios unitarios en  $\varepsilon_{z_{t-1}}$  y $\varepsilon_{y_{t-1}}$  sobre  $y_t$ , respectivamente.
- Los efectos acumulados de los impulsos unitarios en  $\varepsilon_{z_{t-1}}$  y  $\varepsilon_{y_{t-1}}$  se pueden obtener mediante la suma adecuada de los coeficientes de las funciones de impulso respuesta.
- Por ejemplo, después de n períodos, el efecto de  $\varepsilon_{z_{t-1}}$  en el valor de  $y_{t+n}$  es  $\phi_{12}(n)$  Así, después de n períodos, la suma acumulada de los efectos de  $\varepsilon_{z_{t-1}}$  en la secuencia  $y_t$  es

$$\sum_{i=0}^{n} \phi_{12}(i)$$



- ▶ Si  $z_t$  en  $y_t$  son estacionarias, los valores de  $\phi_{ik}(0)$  convergen a cero.
- Los cuatro conjuntos de coeficientes  $\phi_{11}(i)$ ,  $\phi_{12}(i)$ ,  $\phi_{21}(i)$ ,  $\phi_{22}(i)$  denominan funciones impulso respuesta.
- ▶ Si podemos recuperar los parametros del VAR estructural, podemos obtener los errores estructurales  $\varepsilon_{z_t}$  y  $\varepsilon_{y_t}$ , y con ellos los coeficientes  $\phi_{jk}(i)$
- La Descomposición de Choleski permite la identificación de los parámetros estructurales:  $y_t$  no tiene un efecto contemporáneo sobre  $z_t$ . Formalmente, esta restricción se representa estableciendo  $b_{21} = 0$ .

Los términos de error se pueden descomponer de la siguiente manera:

$$e_{1t} = \varepsilon_{yt} - b_{12}\varepsilon_{zt}$$

$$e_{2t} = \varepsilon_{zt}$$

- Se dice que este un "orden" de las variables. Una shock  $\varepsilon_{z_t}$  afecta directamente a  $e_{1t}$  y  $e_{2t}$ , pero una descarga  $\varepsilon_{y_t}$ no afecta a  $e_{2t}$ . Por lo tanto, se dice que  $z_t$  es "causa previamente" a  $y_t$
- La Descomposición de Choleski restringe el sistema: un shock  $\varepsilon_{y_t}$  no tiene efecto directo sobre  $z_t$ , pero hay un efecto indirecto en que los valores rezagados de  $y_t$  afectan el valor contemporáneo de zt.
- La Descomposición fuerza una asimetría potencialmente importante en el sistema ya que un shcok  $\varepsilon_{z_t}$  tiene efectos contemporáneos en yt y zt.

- Suponga que se estima un VAR(1),  $x_t = A_0 + A_1x_{t-1} + e_t$ , y se estima  $A_0$  y  $A_1$ .
- ➤ Si la data disponible hasta la observacion t,. se puede realizar el pronóstico y obtener el error del pronóstico, para los siguientes periodos:

Periodo	Pronóstico	Error pronóstico
t+1:	$E_t x_{t+1} = A_0 + A_1 x_t$	$e_{t+1} = x_{t+1} - E_t(x_{t+1})$
t+2:	$E_t x_{t+2} = A_0 + A_1 A_0 + A_1^2 x_t$	$e_{t+2} = x_{t+2} - E_t(x_{t+2})$
:	<u>:</u>	:
t+n:	$E_t(x_{t+n}) = (I + A_1 + + A_1^{n-1})A_0 + A_1^n x_t$	$e_{t+n} = x_{t+n} - E_t(x_{t+n})$

- Es conveniente escribir los pronosticos en funcion de los errores estructurales.
- ▶ Usando la expresion  $x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$ , podemos obtener el pronostico n periodo, tenemos

$$x_{t+n} = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$$

▶ El error de pronostico en terminos de los errores estructurales

$$x_{t+n} - E_t(x_{t+n}) = \sum_{i=0}^{n-1} \phi_i \varepsilon_{t-i}$$

▶ Enfocandose, por ejemplo, en  $y_t$ , podemos obtener el error de pronostico de n periodos.

$$y_{t+n} - E_t y_{t+n} = \phi_{11}(0)\varepsilon_{yt+n} + \phi_{11}(1)\varepsilon_{yt+n-1} + \dots + \phi_{11}(n-1)\varepsilon_{yt+1} + \phi_{12}(0)\varepsilon_{zt+n} + \phi_{12}(1)\varepsilon_{z+n-1} + \dots + \phi_{12}(n-1)\varepsilon_{zt+1}$$

La varianza del error de pronostico de  $y_{t+n}$  es

$$\sigma_{y}(n)^{2} = \sigma_{y}^{2} \left[ \phi_{11}(0)^{2} + \phi_{11}(1)^{2} + \dots + \phi_{11}(n-1)^{2} \right]$$
$$+ \sigma_{z}^{2} \left[ \phi_{12}(0)^{2} + \phi_{12}(1)^{2} + \dots + \phi_{12}(n-1)^{2} \right]$$

Dado que todos los valores de  $\phi_{jk}(i)^2$  son necesariamente no negativos, la varianza del error de pronostico aumenta a medida que aumenta el horizonte de pronóstico n.

- ► Tenga en cuenta que es posible descomponer la varianza del error de pronóstico de n pasos adelante en las proporciones debidas a cada choque.
- Las proporciones de  $\sigma_y(n)^2$  debido a los choques en las secuencias  $\varepsilon_{yt}$  y  $\varepsilon_{zt}$

$$\frac{\sigma_y^2 \left[\phi_{11}(0)^2 + \phi_{11}(1)^2 + \dots + \phi_{11}(n-1)^2\right]}{\sigma_y(n)^2}$$

$$\frac{\sigma_z^2 \left[\phi_{12}(0)^2 + \phi_{12}(1)^2 + \dots + \phi_{12}(n-1)^2\right]}{\sigma_y(n)^2}$$

- La descomposición de la varianza del error de pronóstico indica la proporción de los movimientos en una secuencia debido a sus "propios" shocks versus los shocks de otra variable.
- Si las perturbaciones de  $\varepsilon_{z_t}$  no explican la varianza del error de pronóstico de  $y_t$  en todos los horizontes de pronóstico, podemos decir que la secuencia  $y_t$  es exógena. En esta circunstancia,  $y_t$  evoluciona independientemente de los shocks  $\varepsilon_{zt}$  y la secuencia  $z_t$ .
- En el otro extremo, las perturbaciones  $\varepsilon_{zt}$  podrían explicar toda la varianza del error de pronóstico en la secuencia  $y_t$  en todos los horizontes de pronóstico, por lo que  $y_t$  sería completamente endógena.
- Es comun que una variable explique casi la totalidad de su varianza del error de pronóstico en horizontes cortos y proporciones más pequeñas en horizontes más largos.



- Tenga en cuenta que la descomposición de la varianza contiene el mismo problema inherente al análisis de la función impulso-respuesta. Para identificar las secuencias  $\varepsilon_{z_t}$  y la  $\varepsilon_{y_t}$  es necesario restringir la matriz B.
- La descomposición de Choleski requiere que toda la varianza del error de pronóstico de  $z_t$  en el período t se deba al  $\varepsilon_{z_t}$ . Si usamos el orden alternativo, toda la varianza del error de pronóstico de un período de  $y_t$  se debería a  $\varepsilon_{y_t}$ .
- Los efectos de estos supuestos alternativos se reducen en horizontes de pronóstico más largos. En la práctica, es útil examinar las descomposiciones de la varianza en varios horizontes de pronóstico. A medida que n aumenta, las descomposiciones de la varianza deberían converger. Además, si el coeficiente de correlación  $\rho_{12}$  es significativamente diferente de cero, es habitual obtener las descomposiciones de la varianza en varios ordenamientos.

## **Granger Causality**

Considere el siguiente VAR(p) con dos variables,

$$y_t = a_{10} + a_{11}^y y_{t-1} + a_{12}^y y_{t-2} + a_{11}^z z_{t-1} + a_{12}^z z_{t-2} + e_{1_t}$$

$$z_t = a_{20} + a_{21}^y y_{t-1} + a_{22}^y y_{t-2} + a_{21}^z z_{t-1} + a_{22}^z z_{t-2} + e_{2_t}$$

- Una prueba de causalidad es si los rezagos de una variable entran en la ecuación de otra variable. En un modelo de dos ecuaciones con 2 rezagos,  $y_t$  no causa Granger  $z_t$  si y solo si todos los coeficientes de  $a_{21}^y = a_{22}^y = 0$ . Por lo tanto, si  $y_t$  no mejora el pronóstico de  $z_t$ , entonces  $y_t$  no causa Granger  $z_t$ .
- Si todas las variables del VAR son estacionarias, la forma directa de probar la causalidad de Granger es usar una prueba F estándar de la restricción:  $a_{21}^{y} = a_{22}^{y} = 0$

## **Granger Causality**

- ▶ Granger Causality es diferente a una prueba de exogeneidad. Para que  $z_t$  sea exógeno, no debe ser afectado por el valor contemporáneo de  $y_t$ .
- Sin embargo, la causalidad de Granger se refiere únicamente a los efectos de los valores pasados de  $y_t$  sobre el valor actual de  $z_t$ . Por tanto, la causalidad de Granger realmente mide si los valores actuales y pasados de  $y_t$  ayudan a pronosticar los valores futuros de  $z_t$ .
- Una ilustración formal, considere el modelo:

$$z_t = \bar{z} + \phi_{21}(0)\varepsilon_{yt} + \sum_{i=0}^{\infty} \phi_{22}(i)\varepsilon_{zt-i}$$

► El pronostico de  $z_{t+1}$  es igual ya sea que este condicionado a  $\varepsilon_{z_{t-i}}$  o para ambos  $\varepsilon_{z_{t-i}}$  y  $\varepsilon_{y_{t-i}}$ ,  $\forall i = 0, 1, ...$  Formalmente:  $E_t(z_{t+1}|z_t) = E_t(z_{t+1}|z_t, y_t)$ 

## Pruebas sobre variables no estacionarias

- 1. Puede utilizar pruebas t o pruebas F en las variables estacionarias.
- Puede realizar una prueba sobre la longitud de rezagos en cualquier variable o conjunto de variables. Esto es cierto independientemente de si la variable en cuestión es estacionaria.
- 3. Es posible que pueda utilizar una prueba F para determinar si una variable no estacionaria Granger causa otra variable no estacionaria. Si se puede hacer que la variable causal aparezca sólo en las primeras diferencias, la prueba está permitida.
- 4. Si el VAR se puede escribir completamente en las primeras diferencias, las pruebas de hipótesis se pueden realizar en cualquier ecuación o cualquier conjunto de ecuaciones utilizando pruebas t o pruebas F. Esto se debe a que todas las variables son estacionarias.

- ▶ El enfoque Sims (1980) sobre VAR tiene la propiedad deseable de que todas las variables se tratan simétricamente, de modo que todas las variables son conjuntamente endógenas y el analista no se basa en restricción de identificación imposible de creer.
- Sin embargo, dada la naturaleza algo ad hoc de la descomposición de Choleski, la belleza del enfoque parece disminuir al construir funciones impulso respuesta y descomposiciones de varianza de error de pronóstico.
- Además, el enfoque VAR ha sido criticado por carecer de contenido económico. El único papel del economista es sugerir las variables apropiadas para incluir en el VAR. A partir de ese momento, el procedimiento es casi mecánico.
- Sin embargo, es posible utilizar la teoría económica para imponer restricciones a las variables para que los resultados no sean ad hoc.



- ▶ El enfoque Sims (1980) sobre VAR tiene la propiedad deseable de que todas las variables se tratan simétricamente, de modo que todas las variables son conjuntamente endógenas y el analista no se basa en restricción de identificación imposible de creer.
- Sin embargo, dada la naturaleza algo ad hoc de la descomposición de Choleski, la belleza del enfoque parece disminuir al construir funciones impulso respuesta y descomposiciones de varianza de error de pronóstico.
- Además, el enfoque VAR ha sido criticado por carecer de contenido económico. El único papel del economista es sugerir las variables apropiadas para incluir en el VAR. A partir de ese momento, el procedimiento es casi mecánico.



- ➤ Sin embargo, es posible utilizar la teoría económica para imponer restricciones a las variables para que los resultados no sean ad hoc.
- ▶ A menos que el modelo estructural subyacente pueda identificarse a partir del modelo VAR de forma standard, las innovaciones en una descomposición de Choleski no tienen una interpretación económica directa.
- ▶ El objetivo de un VAR estructural es utilizar la teoría económica (en lugar de la descomposición de Cholesky) para recuperar las innovaciones estructurales desde los residuos  $e_{1t}$  y  $e_{2t}$ .

La descomposición de Cholesky en realidad hace una fuerte suposición sobre los errores estructurales subyacentes ( $b_{21} = 0$ ). Podemos recuperar los errores estructurales

$$\varepsilon_{yt} = e_{1t} + b_{12}e_{2t}$$
$$\varepsilon_{zt} = e_{2t}$$

- ► Forzar b21 = 0 equivale a suponer que una innovación en yt no tiene un efecto contemporáneo en zt. A menos que exista una base teórica para este supuesto, los choques subyacentes se identifican incorrectamente. Como tal, las respuestas de impulso y las descomposiciones de la varianza que resultan de esta identificación incorrecta pueden ser bastante engañosas.
- Si el coeficiente de correlación entre  $e_{1t}$  y  $e_{2t}$  es bajo, es probable que el orden no sea importante. Sin embargo, en un VAR con varias variables, es improbable que

Facultad de Ciencias sociales y humanísticas

# VAR Estructural: Descomposicion Sims-Bernanke

- ▶ Sims (1986) y Bernanke (1986) propusieron modelar los shocks estructurales utilizando el análisis económico. La idea básica es estimar las relaciones entre los shocks estructurales utilizando un modelo económico.
- Consideremos un VAR(1) de dos variables

$$\left[\begin{array}{cc} 1 & b_{12} \\ 0 & 1 \end{array}\right] \left[\begin{array}{c} y_t \\ z_t \end{array}\right] = \left[\begin{array}{c} b_{10} \\ b_{20} \end{array}\right] + \left[\begin{array}{cc} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{array}\right] \left[\begin{array}{c} y_{t-1} \\ z_{t-1} \end{array}\right] + \left[\begin{array}{c} \varepsilon_{yt} \\ \varepsilon_{zt} \end{array}\right]$$

Matricialmente  $Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t$ . Premultiplicando por  $B^{-1}$ , tenemos  $x_t = A_0 + A_1 x_{t-1} + e_t$ . Donde  $A_0 = B^{-1} \Gamma_0$ ,  $A_1 = B^{-1} \Gamma_1$  y  $e_t = B^{-1} \varepsilon_t$